

APLIKASI KLASIFIKASI ARTIKEL TEKNOLOGI INFORMASI  
PADA MAJALAH CHIP MENGGUNAKAN NAÏVE BAYES

SKRIPSI



Oleh :

RIGTIANTO SETIAWAN ( 0735010056 )

PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNOLOGI INDUSTRI  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JATIM  
2012

**SKRIPSI**  
**APLIKASI KLASIFIKASI ARTIKEL TEKNOLOGI INFORMASI**  
**PADA MAJALAH CHIP MENGGUNAKAN NAÏVE BAYES**

Disusun Oleh :

**RIGTIAN TO SETIAWAN**  
**0735010064**

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Tugas Akhir  
Program Studi Sistem Informasi Fakultas Teknologi Industri  
Universitas Pembangunan Nasional "Veteran" Jawa Timur  
Pada Tanggal 15 Juni 2012  
Telah diperiksa, diuji dan disetujui :

Pembimbing :

1.

Nur Cahyo Wibowo, S.Kom, M.kom  
NPT. 3 7903 04 0197 1

2.

Doddy Ridwandono, S.Kom  
NPT. 278050740218

Tim Penguji :

1.

Prisa Marga Kusumantara, S.Kom  
NPT. 38 211 0602 061

2.

Priza Pandunata, S.Kom  
NPT. 38 301 0602 121

3.

Tikno, ST, M.Kom

Mengetahui,  
Dekan Fakultas Teknologi Industri  
Universitas Pembangunan Nasional "Veteran" Jawa Timur

Ir. Sutiyono, MT.  
NIP. 19600713 198703 1 001

LEMBAR PENGESAHAN

APLIKASI KLASIFIKASI ARTIKEL TEKNOLOGI INFORMASI  
PADA MAJALAH CHIP MENGGUNAKAN NAÏVE BAYES

Disusun oleh :

RIGTIAN TO SETIAWAN  
0735010064

Telah disetujui mengikuti Ujian Negara Lisan  
Periode VI Tahun Akademik 2011/2012

Pembimbing I

Pembimbing II

Nur Cahyo Wibowo, S.Kom, M.Kom  
NPT. 3 7903 04 0197 1

Doddy Ridwandono, S.Kom  
NPT. 278050740218

Mengetahui,  
Ketua Program Studi Sistem Informasi  
Fakultas Teknologi Industri  
Universitas Pembangunan Nasional "Veteran" Jawa Timur

Nur Cahyo Wibowo, S.Kom, M.Kom  
NPT. 3 7903 04 0197 1

YAYASAN KESEJAHTERAAN PENDIDIKAN DAN PERUMAHAN  
UNIVERSITAS PEMBANGUAN NASIONAL “VETERAN” JAWA TIMUR  
FAKULTAS TEKNOLOGI INDUSTRI  
PANITIA UJIAN SKRIPSI / KOMPREHENSIF

---

KETERANGAN REVISI

Mahasiswa di bawah ini :

Nama : RIGTIANTO SETIAWAN

NPM : 0735010064

Program Studi : Sistem Informasi

Telah mengerjakan revisi skripsi dengan judul :

APLIKASI KLASIFIKASI ARTIKEL TEKNOLOGI INFORMASI PADA  
MAJALAH CHIP MENGGUNAKAN NAÏVE BAYES

Oleh karenanya mahasiswa tersebut diatas dinyatakan bebas revisi skripsi dan diijinkan untuk membukukan skripsi dengan judul tersebut.

Surabaya, 22 Juni 2012

Dosen Penguji yang memerintahkan revisi:

- |   |   |  |
|---|---|--|
| 1.) <u>Prisa Marga Kusumantara, S.Kom</u> | { |  |
| NPT. 3 7903 04 0197 1                     |   |  |
| 2.) <u>Priza Pandunata, S.Kom</u>         | { |  |
| NPT. 38 301 0602 121                      |   |  |
| 3.) <u>Tikno, ST, M.Kom</u>               | { |  |

Mengetahui,  
Dosen Pembimbing I

Mengetahui,  
Dosen Pembimbing II

Nur Cahyo Wibowo, S.Kom, M.Kom  
NPT. 3 7903 04 0197 1

Doddy Ridwandono, S.Kom  
NPT. 278050740218

## KATA PENGANTAR

Syukur Alhamdulillah, puji dan syukur kami panjatkan kehadiran Allah SWT, karena dengan limpah dan rahmat dan karunia-nya penulis dapat menyelesaikan penyusunan laporan tugas akhir “APLIKASI KLASIFIKASI ARTIKEL TEKNOLOGI INFORMASI PADA MAJALAH CHIP MENGGUNAKAN NAÏVE BAYES”.

Maksud dan tujuan penyusunan laporan tugas akhir adalah untuk memenuhi salah satu syarat dalam perkuliahan guna pembelajaran mahasiswa di dunia kerja yang akan di implementasikan sepenuhnya berdasarkan dengan keahlian mahasiswa tersebut. Dan untuk mencapai gelar sarjana pada Fakultas Teknologi Industri, Jurusan Sistem Informasi Pembangunan Nasional “Veteran” Jawa Timur.

Pada kesempatan ini penulis menyampaikan terima kasih kepada semua pihak yang telah membantu penulis dalam menyelesaikan laporan ini. Secara khusus penulis menyampaikan terima kasih kepada :

1. Tugas akhir ini penulis dedikasikan kepada kedua orangtua penulis, yaitu Bapak Soeprapto dan Ibu Erny yang telah sabar dan mencurahkan segala kasih sayangnya kepada penulis, serta dukungan material dan spiritual selama masa studi di UPN. Kedua orang tua yang telah memberi dukungan doa dan materi.
2. Bapak Nur Cahyo Wibowo S.Kom, M. Kom. Selaku dosen pembimbing I.
3. Bapak Novyanto Puji Raharjo S.Kom. Selaku dosen pembimbing II.
4. Seluruh dosen Universitas Pembangunan Nasional ”Veteran” Jawa Timur.

5. Semua teman-teman yang sudah membantu proses penyelesaian laporan ini khususnya seperti: Bayu Topan, Aprianto, Dhiparendra, Edi Saktia, Adhe Perdana mahasiswa Sistem Informasi Angkatan 2007.
6. Alm. Yanuar Rahadi. Sahabat yang berjuang bersama saat SP dan UK. Teman, aku sudah lulus sekarang. Semoga amal dan ibadahmu diterima disisi-Nya.
7. Teman Ruri Pamungkas dan Muhamad Yunus Reynaldi yang berada di Jakarta yang sudah memberi dukungan menyelesaikan tugas akhir ini.
8. Kepada teman teman Administrasi Negara UPN angkatan 2007 dan 2008, terima kasih telah menemani dan menghibur dikala galau skripsi, terutama terimakasih kepada Muchamad Abie Firdaus yang mengajari banyak hal.
9. SANTOSO (Aprilia, Arta, Tata dan koko Doni yang sudah menemani di Surabaya di saat sendiri.
10. Wasono Setyo Prayogo dan Ambar Sulisetyowati yang sudah membimbing saya selama ini.
11. Semua pihak yang telah membantu proses penyelesaian laporan ini yang tidak mungkin penulis sebutkan satu persatu.

Dengan keterbatasan ilmu dan pengalaman, penulis dengan rendah hati menyadari sepenuhnya, bahwa penyusunan laporan ini masih jauh dari sempurna, oleh karena itu penulis bersedia dan ikhlas lahir bathin untuk menerima segala saran dan kritik dari berbagai pihak demi perbaikan laporan ini.

Surabaya, 31 Mei 2012

Penyusun

## DAFTAR ISI

ABSTRAKSI .....	i
KATA PENGANTAR .....	ii
DAFTAR ISI .....	iv
DAFTAR TABEL .....	vii
DAFTAR GAMBAR .....	viii
BAB I PENDAHULUAN	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	1
1.3 Batasan Masalah .....	2
1.4 Tujuan .....	2
1.5 Manfaat .....	3
1.6 Sistematika Penulisan .....	3
BAB II TINJAUAN PUSTAKA	
2.1 Pembelajaran Mesin.....	5
2.1.1 Komponen Sistem Cerdas .....	7
2.1.2 Rekayasa Pengetahuan .....	10
2.2. Proses Pengklasifikasian .....	11
2.3. Metode Klasifikasi Teks .....	11
2.4 Representasi Dokumen .....	12
2.5 Mprfologi Bahasa Indonesia .....	12
2.6 Pembobotan Kata .....	14
2.7 Training Set Dan Testing Set .....	16

2.8	Metode Naïve Bayes .....	16
2.9	Karakteristik Naïve Bayes Classifier .....	19
2.10	Tahapan Pra Proses .....	22
2.11	Text Mining .....	25
2.12	Diagram Proses Klasifikasi .....	27
2.13	Delphi .....	28
2.14	Majalah Chips .....	38

### BAB III ANALISIS DAN PERANCANGAN

3.1	Desain Penelitian .....	41
3.2	Bahan Penelitian .....	42
3.3	Structured Chart Sistem Klasifikasi Artikel .....	43
3.4	Diagram Work Flow .....	44
3.5	Diagram Alir Sistem .....	45
3.6	Perancangan Antar Muka .....	46

### BAB IV HASIL DAN PEMBAHASAN

4.1	Perangkat Keras Yang Dipergunakan .....	49
4.2	Perangkat Lunak Yang Dipergunakan .....	49
4.3	Implementasi Prosedur Proses Klasifikasi .....	50
4.4	Implementasi Prosedur Proses Parsing .....	51
4.5	Implementasi Prosedur Proses Filtering .....	52
4.6	Implementasi Prosedur Proses Stemming .....	53
4.7	Implementasi Prosedur Bayes Classifier .....	54
4.8	Data Pelatihan .....	57
4.9	Data Pengujian .....	60



4.9	Hasil Pengujian .....	62
BAB V PENUTUP		
5.1	Kesimpulan .....	71
5.2	Saran .....	71
DAFTAR PUSTAKA .....		72

## DAFTAR TABEL

Tabel 2.1	Contoh Frekuensi Kata dalam Suatu Dokumen .....	14
Tabel 2.2	Bobot Kata Dalam Vektor Dokumen .....	15
Tabel 2.3	Bobot Kata Setelah Proses Normalisasi .....	16
Tabel 2.4	Data Training .....	17
Tabel 4.1	Spesifikasi Perangkat Komputer .....	49
Tabel 4.2	Spesifikasi Perangkat Lunak .....	49
Tabel 4.3	Hasil Pengujian Probabilistic Model .....	59
Tabel 4.4	Hasil Perhitungan Bayes Pada File Uji_1.txt .....	63
Tabel 4.5	Hasil Perhitungan Bayes Pada File Uji_2.txt .....	64

## DAFTAR GAMBAR

Gambar 2.1	Skema Sistem Pakar .....	7
Gambar 2.2	Proses Rekayasa Pengetahuan .....	10
Gambar 2.3	Tahapan Pra Proses .....	24
Gambar 2.4	Tahapan Umum Text Mining .....	12
Gambar 2.5	Contoh Tokenisasi .....	26
Gambar 2.6	Contoh Filtering .....	26
Gambar 2.7	Contoh Stemmisasi .....	27
Gambar 2.8	Diagram Proses Klasifikasi .....	27
Gambar 3.1	Skema Desain Penelitian .....	41
Gambar 3.2	Structured Chart Klasifikasi Artikel .....	43
Gambar 3.3	Diagram Workflow Sistem Klasifikasi Artikel .....	44
Gambar 3.4	Diagram Alir Sistem .....	45
Gambar 3.5	Rancangan Form Utama .....	46
Gambar 3.6	Rancangan Form Training .....	47
Gambar 3.7	Rancangan Form Stemming .....	48
Gambar 3.8	Rancangan Form Stoplist .....	48
Gambar 4.1	Hasil Pengujian artikel uji_1.txt .....	63
Gambar 4.2	Hasil Pengujian Dengan Menggunakan Artikel uji_2.txt..	64
Gambar 4.3.	Hasil Pengujian Dengan Menggunakan Artikel Dell.txt .....	65
Gambar 4.4.	Hasil Pengujian Dengan Menggunakan Artikel Amd.txt .....	65
Gambar 4.5.	Hasil Pengujian Dengan Menggunakan Artikel Android.txt..	66
Gambar 4.6.	Hasil Pengujian Dengan Menggunakan Artikel Intel.txt .....	66

Gambar 4.7. Hasil Pengujian Dengan Menggunakan Artikel Apel.txt .....	67
Gambar 4.8. Hasil Pengujian Dengan Menggunakan Artikel Notebook.txt.	67
Gambar 4.9. Hasil Pengujian Dengan Menggunakan Artikel Dunia Gaming.txt .....	68
Gambar 4.10. Hasil Pengujian Dengan Menggunakan Artikel Backup Data.txt .....	68
Gambar 4.11. Hasil Pengujian Dengan Menggunakan Artikel Tips Blackberry.txt .....	69
Gambar 4.12. Hasil Pengujian Dengan Menggunakan Artikel Marhaban.txt .....	69

Judul : APLIKASI KLASIFIKASI ARTIKEL TEKNOLOGI INFORMASI  
PADA MAJALAH CHIP MENGGUNAKAN NAÏVE BAYES

Dosen 1 : Nur Cahyo Wibowo, S.Kom, M.Kom  
Dosen 2 : Doddy Ridwandono, S.Kom  
Penulis : Rigtianto Setiawan

---

### Abstrak

Pengklasifikasian teks sangat dibutuhkan dalam berbagai macam aplikasi, terutama aplikasi yang jumlah dokumennya bertambah dengan cepat seiring dengan bertambahnya waktu. Contohnya adalah aplikasi spam email, indek otomatis pada artikel ilmiah, dan lain sebagainya. Ada dua varian utama dalam penggolongan teks, yaitu clustering text dan klasifikasi teks

Permasalahan yang dihadapi adalah bagaimana membuat aplikasi rekomendasi pengklasifikasian suatu artikel teks pada CHIP menggunakan algoritma Naïve Bayes dan seberapa akuratkah proses klasifikasi dalam mengklasifikasikan artikel teks majalah CHIP menggunakan algoritma Naïve Bayes.

Pembuatan aplikasi ini dikembangkan menggunakan bahasa pemrograman Delphi dan diintegrasikan dengan database Microsoft Access. Aplikasi yang dihasilkan dapat mengklasifikasikan artikel teks pada majalah komputer Chip dengan baik.

Kata kunci: Klasifikasi Teks, Naïve Bayes, Majalah Chip.

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Pengklasifikasian teks sangat dibutuhkan dalam berbagai macam aplikasi, terutama aplikasi yang jumlah dokumennya bertambah dengan cepat seiring dengan bertambahnya waktu. Contohnya adalah aplikasi spam email, indek otomatis pada artikel ilmiah, dan lain sebagainya. Ada dua varian utama dalam penggolongan teks, yaitu clustering text dan klasifikasi teks. Clustering teks berhubungan dengan menemukan sebuah struktur kelompok yang belum kelihatan (tak terpandu atau unsupervised) dari sekumpulan dokumen. Sedangkan pengklasifikasian teks dapat dianggap sebagai proses untuk membentuk golongan-golongan (kelas-kelas) dari dokumen berdasarkan pada kelas kelompok yang sudah diketahui sebelumnya (terpandu atau supervised). Naive Bayes merupakan salah satu metode yang digunakan dalam pengklasifikasian teks. Metode lainnya adalah k-Nearest Neighbor, Support Vector, Rochio Classifier dan lain-lain.

### 1.2. Perumusan Masalah

Rumusan masalah yang digunakan dalam tugas akhir ini adalah :

- a) Bagaimana membuat aplikasi rekomendasi pengklasifikasian suatu artikel teks pada majalah CHIP menggunakan algoritma naive bayes
- b) Seberapa akuratkah proses klasifikasi dalam mengklasifikasikan artikel teks majalah Chip dengan menggunakan algoritma naive bayes ?

### 1.3. Batasan Masalah

Dalam pembuatan aplikasi teks untuk tugas akhir ini, menggunakan batasan masalah sebagai berikut :

- a) Artikel teks yang dipergunakan tidak lebih dari 150 kata
- b) Artikel teks yang dimaksud berupa file dengan ekstensi \*.txt
- c) Artikel yang dipergunakan sebagai pengujian diambil dari majalah komputer Chips
- d) Pengklasifikasian artikel berdasarkan kategori yang telah ada di majalah Chips dan ditambah satu yaitu tidak terkategori.
- e) Kategori setiap artikel yang akan diklasifikasikan telah ditentukan sebelumnya secara manual, hal ini diperlukan untuk pembuatan data pelatihan dan untuk proses klasifikasi. Pengklasifikasian secara manual ini dapat dilakukan oleh peneliti. Nantinya, sistem akan mengecek hasil klasifikasi secara manual dengan hasil klasifikasi secara otomatis menggunakan sistem untuk mendapatkan angka akurasi pengklasifikasian

### 1.4. Tujuan

Tujuan yang ingin dicapai pada pengerjaan tugas akhir ini adalah:

- a) Dapat membuat aplikasi rekomendasi pengklasifikasian artikel teks pada majalah komputer chips.
- b) Dapat menggunakan algoritma Naive Bayes untuk klasifikasi teks.

### 1.5. Manfaat

Adapun manfaat dan tujuan yang ingin diperoleh dari pengerjaan tugas akhir ini adalah :

- a) Mendapat kemampuan pemahaman dan analisa tentang algoritma naive bayes.
- b) Dapat mengimplementasikan algoritma naive bayes untuk membuat aplikasi klasifikasi data dokumen teks.

### 1.6. Sistematika Penulisan

Sistematika penulisan tugas akhir ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut :



**BAB I            PENDAHULUAN**

Bab ini berisi latar belakang masalah, identifikasi masalah, maksud dan tujuan yang ingin dicapai, batasan masalah, metodologi penelitian yang diterapkan dalam memperoleh dan mengumpulkan data, waktu dan tempat penelitian, serta sistematika penulisan.

**BAB II            TINJAUAN PUSTAKA**

Membahas berbagai konsep dasar dan teori-teori yang berkaitan dengan topik masalah yang diambil dan hal-hal yang berguna dalam proses analisis permasalahan.

**BAB III           ANALISIS DAN PERANCANGAN**

Menganalisis masalah dari model penelitian untuk memperlihatkan keterkaitan antar variabel yang diteliti serta model matematis untuk analisisnya.

**BAB IV           HASIL DAN PEMBAHASAN**

Membahas mengenai pengimplementasian aplikasi yang telah dibuat ke perangkat yang akan digunakan serta melakukan pengujian terhadap aplikasi yang telah diimplementasikan tersebut.

**BAB V            PENUTUP**

Berisi kesimpulan dan saran yang sudah diperoleh dari hasil penulisan tugas akhir dan daftar pustaka.